Supplementary information

The emergence of macroscopic interactions between intersecting pedestrian streams

Published in *Transportation Research Part B*

Authors: Nikolai W.F. Bode, Mohcine Chraibi, Stefan Holl

Supplementary Methods

Method for computing density and specific flow

We use the method by Holl [Holl, 2016] to compute the instantaneous pedestrian density in and specific flow through the crossing. This method is a generalisation of Edie's method [Edie, 1965]. We measure density and specific flow in consecutive time intervals of length $\Delta t = 2$ seconds. In the following, we describe how specific flow and density are computed for a single time interval. Supplementary table S1 defines all quantities needed.

Symbol	Description
A	Size of measurement area. The only constraint for the measurement area is that it takes a convex form.
t_0, t_1	Lower and upper temporal limit of the measurement interval, such that $\Delta t = t_1 - t_0 = 2$ seconds.
$t_{i,in}, t_{i,out}$	Time points at which person i enters and exits the measurement area. Note that these time points may occur before, after or within the measurement interval.
$\boldsymbol{x_i}(t)$	Position vector of person i at time t.
$\boldsymbol{a_i}, a_i$	The vector between the entrance and exit points of person i: $\boldsymbol{a_i} = \boldsymbol{x_i}(t_{i,out}) - \boldsymbol{x_i}(t_{i,in})$. The length of this vector is denoted by a_i. This is the shortest route along the main direction of movement for person i.
b_i	Straight line approximation of the path covered by person i inside the measurement area during the measurement interval. Five cases have to be considered: Case 1: if $t_0 \leq t_{i,in}$ and $t_1 > t_{i,in}$ and $t_1 < t_{i,out}$, then $b_i = \lvert\boldsymbol{x_i}(t_{i,in}) - \boldsymbol{x_i}(t_1)\rvert$ Case 2: if $t_0 > t_{i,in}$ and $t_1 < t_{i,out}$, then $b_i = \lvert\boldsymbol{x_i}(t_0) - \boldsymbol{x_i}(t_1)\rvert$ Case 3: if $t_0 > t_{i,in}$ and $t_0 < t_{i,out}$ and $t_1 \geq t_{i,out}$, then $b_i = \lvert\boldsymbol{x_i}(t_0) - \boldsymbol{x_i}(t_{i,out})\rvert$ Case 4: if $t_0 \leq t_{i,in}$ and $t_1 \geq t_{i,out}$, then $b_i = \lvert\boldsymbol{x_i}(t_{i,in}) - \boldsymbol{x_i}(t_{i,out})\rvert$

	Case 5: if $t_1 \leq t_{i,in}$ or $t_0 \geq t_{i,out}$, then $b_i = 0$
c_i	Straight line approximation of the path covered by person i inside the measurement area, but not during the measurement interval. Again, the same five case as in the definition for b_i have to be considered: Case 1: $c_i = \lvert \boldsymbol{x_i}(t_1) - \boldsymbol{x_i}(t_{i,out}) \rvert$ Case 2: $c_i = \lvert \boldsymbol{x_i}(t_{i,in}) - \boldsymbol{x_i}(t_0) \rvert + \lvert \boldsymbol{x_i}(t_1) - \boldsymbol{x_i}(t_{i,out}) \rvert$ Case 3: $c_i = \lvert \boldsymbol{x_i}(t_{i,in}) - \boldsymbol{x_i}(t_0) \rvert$ Case 4: $c_i = 0$ Case 5: $c_i = \lvert \boldsymbol{x_i}(t_{i,in}) - \boldsymbol{x_i}(t_{i,out}) \rvert$
d_i	Dimensionless relationship between the path lengths covered inside the measurement area during and outside of the measurement interval. $d_i = \frac{b_i}{b_i + c_i} \in [0,1]$.
e_i	Length of the distance person i covers along her main movement direction during the measurement interval: $e_i = d_i \times a_i$.
Δt_i	Time person i spends inside the measurement area during the measurement interval, $[t_0, t_1]$.

Table S1: notation and definition for all quantities needed in determining the instantaneous flow and density.

Using these definitions, Holl [Holl, 2016] obtains the mean density over the measurement interval inside the measurement area (where n is the total number of pedestrians):

$$\rho = \frac{1}{A} \times \sum_{i=1}^{n} \frac{\Delta t_i}{\Delta t} \qquad (S1)$$

Similarly, the mean speed along the main movement directions inside the measurement area over the measurement interval is:

$$v = \frac{\sum_{i=1}^{n} e_i}{\sum_{i=1}^{n} \Delta t_i} \qquad (S2)$$

Assuming that the specific flow J_s can be computed as a product of density and speed and using equations S1 and S2, we obtain:

$$J_s = \rho \times v = \frac{1}{\Delta t} \times \frac{1}{A} \times \sum_{i=1}^{n} e_i \qquad (S3)$$

We compute the density and specific flow according to the equations above for consecutive measurement intervals and plot the measurements of specific flow against the density to obtain the fundamental diagram in figure 1 in the main text.

Models for four pedestrian streams

The models for two pedestrian streams presented in the main text can be extend to any number of intersecting streams. Here, we specify the structure of models for four intersecting streams. Recall that we model the number of pedestrians, or population, from each stream inside the crossing. The models take the general form:

$$\frac{dX_1}{dt} = f_{in}(X_1, X_2, X_3, X_4) - f_{out}(X_1, X_2, X_3, X_4)$$

$$\frac{dX_2}{dt} = f_{in}(X_2, X_1, X_3, X_4) - f_{out}(X_2, X_1, X_3, X_4)$$
$$\frac{dX_3}{dt} = f_{in}(X_3, X_1, X_2, X_4) - f_{out}(X_3, X_1, X_2, X_4)$$
$$\frac{dX_4}{dt} = f_{in}(X_4, X_1, X_2, X_3) - f_{out}(X_4, X_1, X_2, X_3) \qquad (S4)$$

Where f_{in} and f_{out} are functions that take positive values and capture the rate of in- and outflow for the crossing, respectively. To complete the model specification, we detail the functional form of the in- and outflow rates.

Model 1 assumes that there are no interactions between the stream populations. The in- and outflow functions for X_i are as follows:

$$f_{in}(X_i, X_j, X_k, X_l) = \frac{\alpha}{[1 + \exp(X_i - \gamma)]}$$
$$f_{out}(X_i, X_j, X_k, X_l) = \mu X_i \exp(-\varepsilon X_i), \qquad (S5)$$

where $(i, j, k, l) \in \{1,2,3,4\}$ and $i \neq j \neq k \neq l$. The interpretation of the model parameters remains the same.

Model 2 introduces interactions between streams via the overall number of pedestrians inside the intersection:

$$f_{in}(X_i, X_j, X_k, X_l) = \frac{\alpha}{[1 + \exp(X_i + X_j + X_k + X_l - \gamma)]}$$
$$f_{out}(X_i, X_j, X_k, X_l) = \mu X_i \exp\big(-\varepsilon[X_i + X_j + X_k + X_l]\big) \qquad (S6)$$

Model 3 implements interactions between the stream populations in a way that is analogous to the case for two intersecting streams in equation 4 in the main text:

$$f_{in}(X_i, X_j, X_k, X_l) = \frac{\alpha}{[1 + \exp(X_i + (X_i X_j X_k X_l)^{\frac{1}{4}} - \gamma)]}$$
$$f_{out}(X_i, X_j, X_k, X_l) = \mu X_i \exp\left(-\varepsilon X_i - \delta (X_i X_j X_k X_l)^{\frac{1}{4}}\right) \qquad (S7)$$

Note how in contrast to equation 4 in the main text, in equation S7, the square-root has been replaced with a quartic root, as we now consider the product of four population sizes.

Stability analysis in the case of four intersecting streams

As it is difficult to visualise the results of a mathematical stability analysis of our models for four intersecting streams graphically, we only discuss the key features arising from this analysis here.

For model 1, we note that equation S5 is identical to equation 2 in the main text, as the dynamics of streams are independent. The findings from our analysis for two intersecting streams therefore extend to the case of four intersecting streams.

For model 2, equating f_{in} and f_{out} using equation S6 for each stream separately and then combining the resulting equations yields the same result as for two intersecting streams, namely that at equilibrium we have $X_1 = X_2 = X_3 = X_4$. Thus, the bifurcation diagram for four intersecting streams will be qualitatively the same as the diagram for two intersecting streams.

Using a similar argument as in the main text, we can find equilibrium points for model 3 in addition to $X_1 = X_2 = X_3 = X_4$ by numerically finding $\{X_1, X_2, X_3, X_4\}$ that satisfy:

$$\frac{\exp(\varepsilon X_i)}{X_i[1+\exp(X_i+(X_iX_jX_kX_l)^{\frac{1}{4}}-\gamma)]}=\frac{\exp(\varepsilon X_j)}{X_j[1+\exp(X_j+(X_iX_jX_kX_l)^{\frac{1}{4}}-\gamma)]} \qquad (S8)$$

$\forall i,j$, where $(i,j,k,l)\in\{1,2,3,4\}$ and $i\neq j\neq k\neq l$. The main difference between models 2 and 3 we would like to show is that in contrast to model 2, model 3 can have equilibrium points for which $X_i\neq X_j$ for some i,j. To show this it is not necessary to perform the complete stability analysis (although equation S8 shows that this is possible). It is sufficient to show that at least one such equilibrium point exists. We do so using simulations (see supplementary figure S5).

We use the same model fitting approach for four intersecting streams as described in the main text.

Assessing the assumption of simulations using the Gillespie algorithm

As already indicated in the main text, our simulations using the Gillespie algorithm assume that the time gaps between events (pedestrians entering or leaving the crossing) follow exponential distributions. According to this, the parameter of the exponential distribution that a given observed time gap is assumed to follow is the sum of the f_{in} and f_{out} rates across all streams predicted by our models [Gillespie, 1977]. As the f_{in} and f_{out} rates depend on the number of pedestrians inside the crossing, the observed time gaps between events do not follow one exponential distribution, but many different exponential distributions. To assess our assumption, we thus make use of the probability integral transform. This method uses the cumulative distribution function of random variables to convert random variables of any distribution into random variables that are uniformly distributed between 0 and 1 (e.g. [Dodge, 2003]). We apply this method to all observed time gaps in our data and expect to find a uniform distribution between 0 and 1 if our assumption holds unequivocally. Where multiple events occur in one time step of our trajectory data, we use an equal temporal spacing and a random order of events. As the obtained transformed time gap distributions vary considerably across the parameter values included in posterior distributions, we report three results: a best-case scenario, a worst-case scenario and a scenario where we sample from the posterior distribution of models and parameters (supplementary figure S6).

Our results show that while for the best-case scenario, the distribution of transformed time gaps is close to a uniform distribution (supplementary figure S6A), it deviates substantially from the expected uniform distribution for the other two scenarios (supplementary figure S6C,D). For these scenarios, the distribution of transformed time gaps has a large peak for values close to 0.

These findings suggest that the validity of the distributional assumption underlying our model is questionable, but they do not invalidate our analysis. First, the best-case scenario suggests that for some parameter choices the model assumption is appropriate. Second, a possible explanation for the peak observed in supplementary figure S6C,D is that the average rate of events in our model may often be under-estimated. This could arise from the fact that we do not fit our models directly to the time gaps between individual events, but only consider aggregated event counts (equation 5 in the main text). Third, our approach to deal with the limited temporal resolution of our data (16 or 25 frames per second) could contribute to the peak in transformed time gaps close to zero. For exponential distributions with large rates, small changes in observed time gaps could have a large effect on the transformed time gaps. Fourth, it is important to note that the stability properties of our models do not depend on the distributional assumption of the Gillespie algorithm and as we make the same assumption for all models in our model fitting approach, the model selection process is fair, and its results hold.

In summary, based on the results in supplementary figure S6, we cannot claim that the distributional assumption we make in simulating our model holds unequivocally. We suggest possible reasons for the results we observe and argue that our main model selection results are valid. To further support this notion and to illustrate the fit of our models to the experimental data, we compare simulations to average pedestrian numbers in streams over time (figure 4 in the main text).

Supplementary References

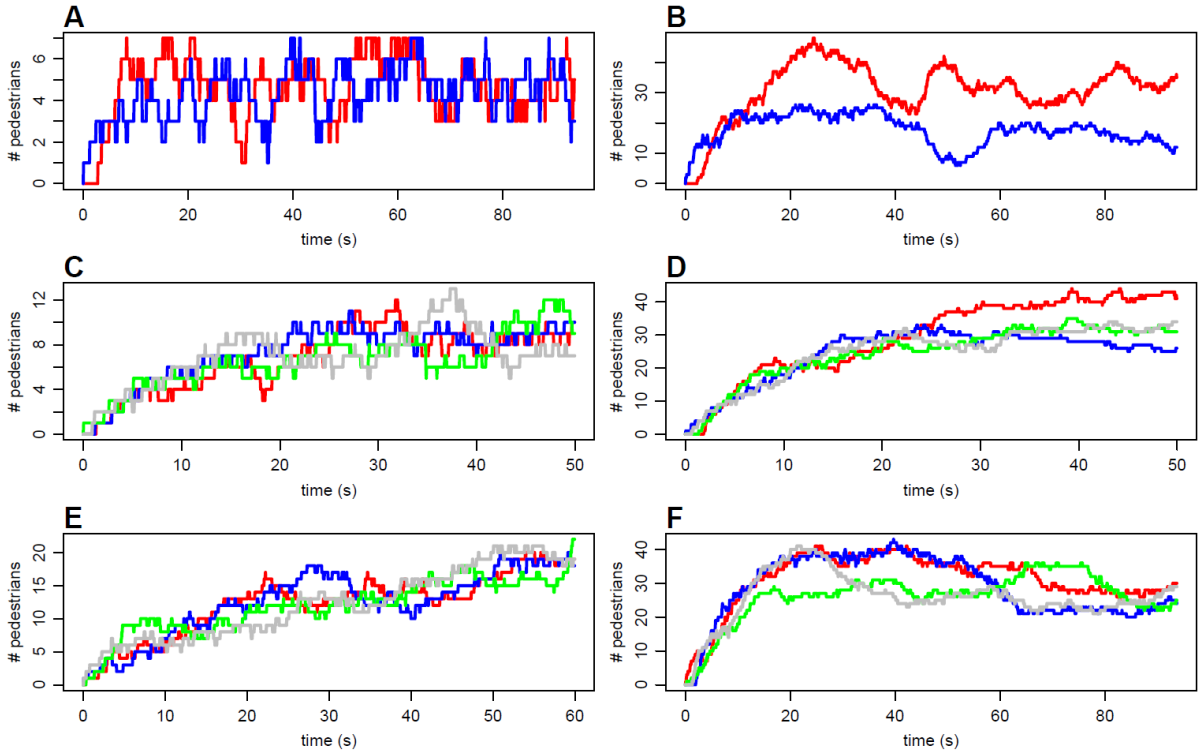
Dodge, Y. (2003) The Oxford Dictionary of Statistical Terms. Oxford University Press, Oxford.

Edie, L.C. (1965). Discussion of traffic stream measurements and definitions. Proceedings of the second international symposium on the theory of traffic flow, London 1963, pp. 139-154.

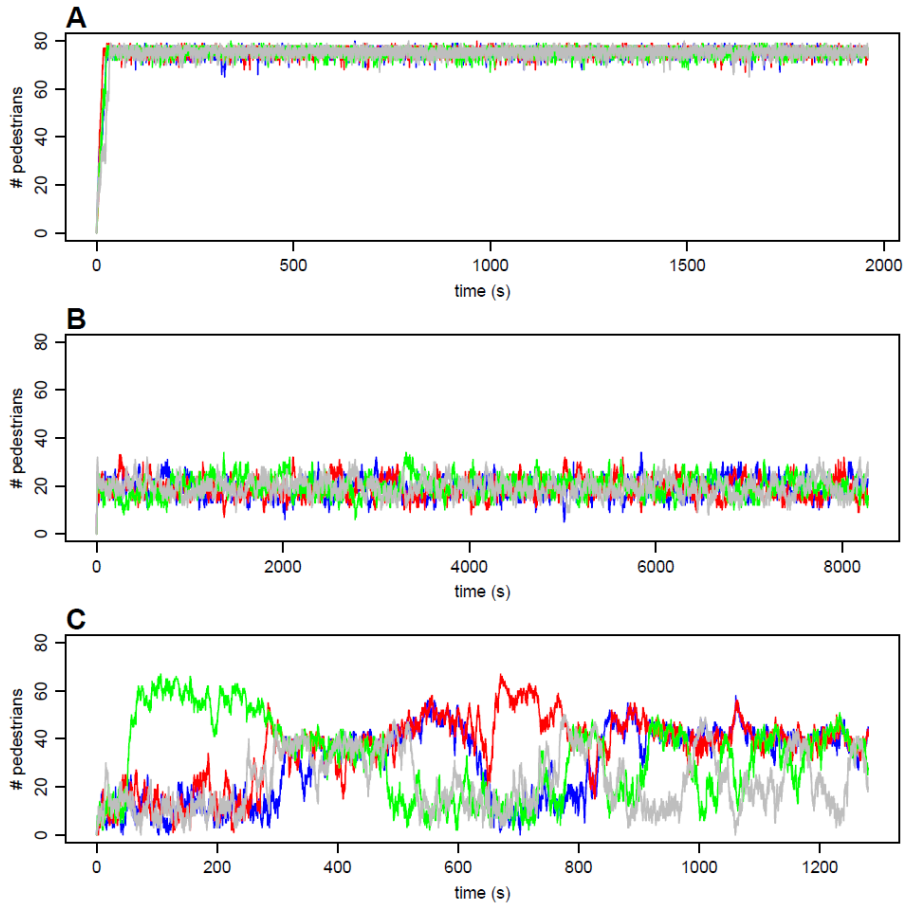
Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. The Journal of Physical Chemistry, 81(25), 2340-2361.

Holl, S. (2016). Methoden für die Bemessung der Leistungsfähigkeit multidirektional genutzter Fußverkehrsanlagen (Doctoral dissertation, Universität Wuppertal, Fakultät für Architektur und Bauingenieurwesen» Bauingenieurwesen» Dissertationen).

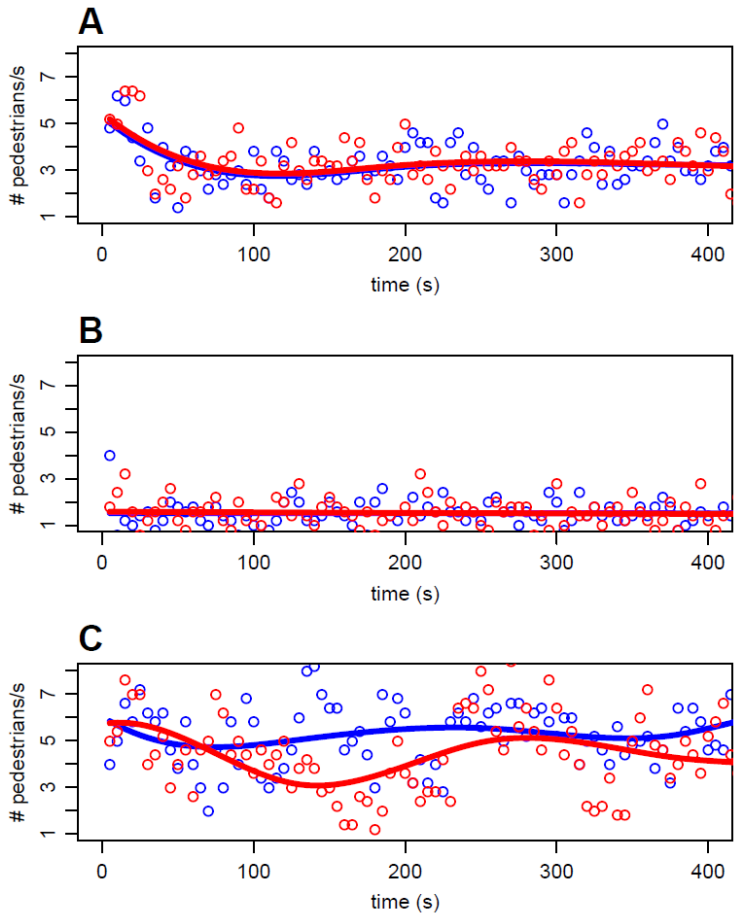
Supplementary figures



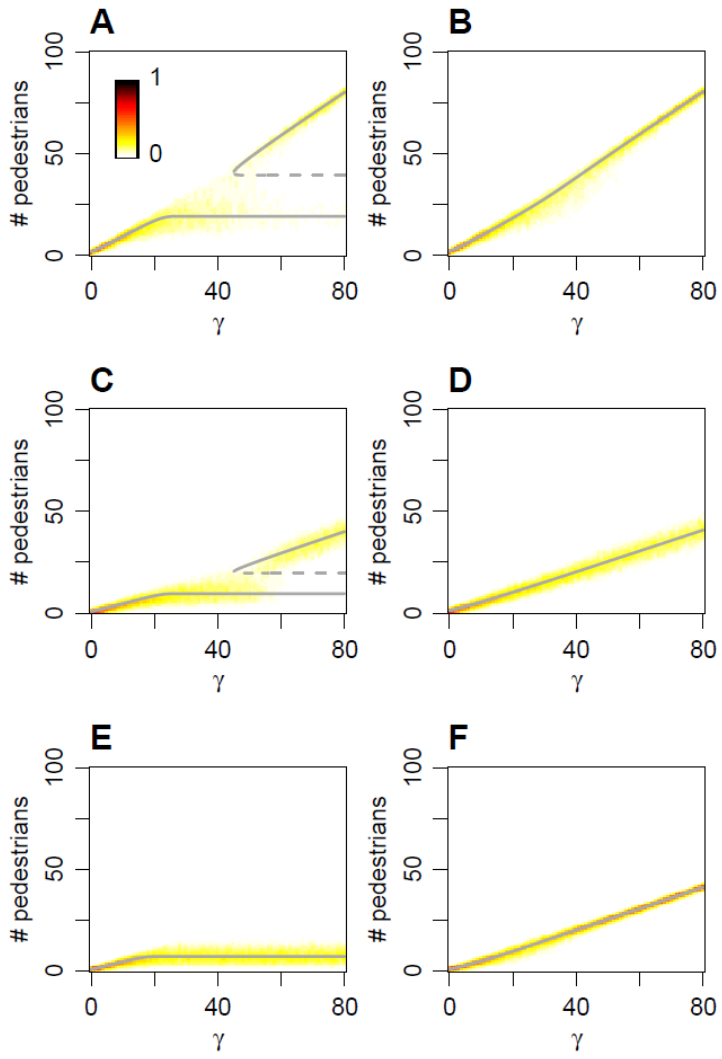
Supplementary figure S1: Time series of pedestrian numbers inside the crossing from experiments. (A,B) show data from experiments with two intersecting streams under the free condition and inflow widths of 0.6 m and 4.0 m, respectively. (C,D) show data from experiments with four intersecting streams under the free condition and inflow widths of 0.6 m and 1.5 m, respectively. (E,F) show data from experiments with four intersecting streams under the column condition and inflow widths of 0.9 m and 4.0 m, respectively. Two streams are shown in blue and red and four streams in blue, red, green and grey.



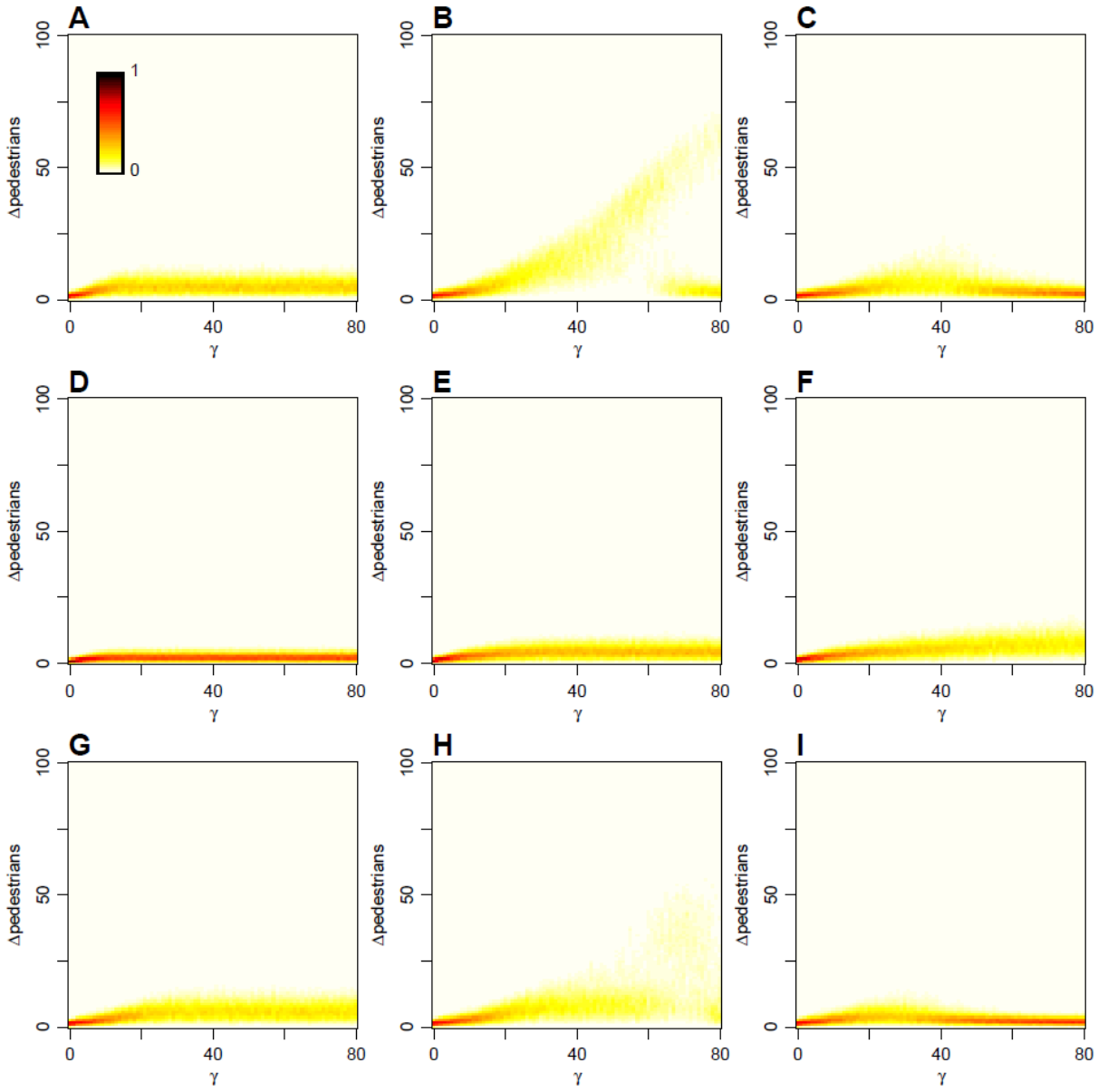
Supplementary figure S2: Example simulations for four intersecting pedestrian streams shown in red, blue, green and grey for model 1 (A), model 2 (B) and model 3 (C). Model parameters are the same as in figure 2 in the main text: α=8, γ=50, ε=0.036 and μ=0.62 for models 1 and 2. For model 3, α=6, γ=50, μ=1.2, ε=0.053 and δ=0.01. For the simulation of model 1, the stream populations are independent of each other, but all take the same value (A). The simulation of model 2 shows that stream populations fluctuate around a common value (B). In the simulation of model 3, switch between several different levels and not all stream populations are necessarily at the same level. All plots are from simulations over 50,000 update steps.



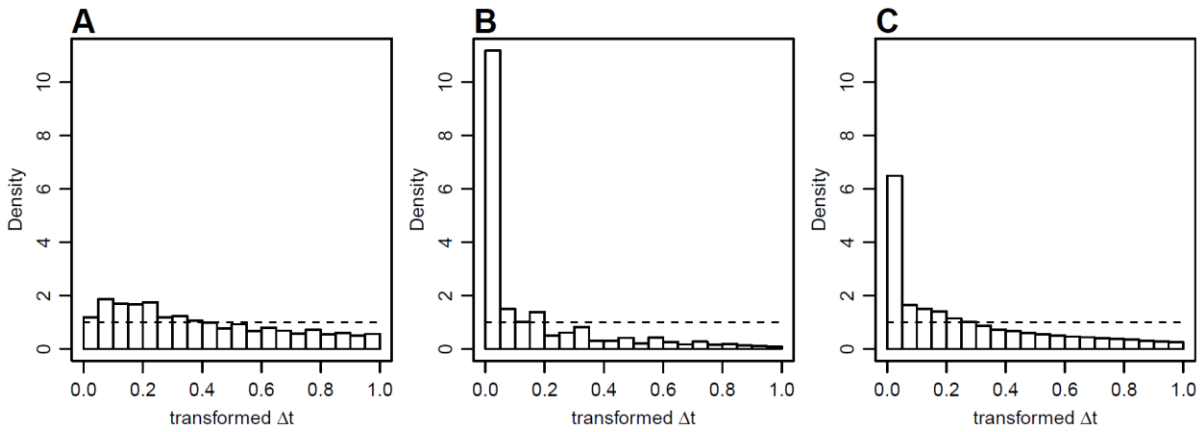
Supplementary figure S3: Outflow of pedestrians for simulations in figure 2B-D in the main text. The two streams are shown in blue and red. The outflow is computed by counting the number of simulated pedestrians leaving the crossing over 5 s intervals. The continuous lines are spline regressions through the data points. Temporal variation in the flow is weak, if present at all, in (A-B), but visible in (C).



Supplementary figure S4: Stability analysis for pedestrian stream population models. Bifurcation diagrams are for the case of two intersecting pedestrian streams for models 1 (A-B), 2 (C-D) and 3 (E-F). Grey lines indicate equilibrium solutions for which $X_1=X_2$ (except for model 1, where X_1 and X_2 are independent). Blue lines show equilibrium solutions for which $X_1 \neq X_2$. Solid lines denote stable and dashed lines unstable equilibrium solutions. Heatmaps underlying the bifurcation diagrams show stream populations observed in the last 100 updates of 100 replicate stochastic simulations over 10,000 simulation steps and show probability densities separately for each value of γ (colour scale in panel A). Parameter values used are $\mu=0.62$ and $\varepsilon=0.04$ for models 1 and 2. For model 3, we use $\alpha=6$, $\mu=1.2$ and $\delta=0.01$. In addition: (A) $\alpha=6$; (B) $\alpha=7$; (C) $\alpha=3$; (D) $\alpha=8$; (E) $\varepsilon=0.04$; (F) $\varepsilon=0.1$.



Supplementary figure S5: Stability analysis for pedestrian stream population models for four intersecting streams. Heatmaps show the distribution of the maximal difference, denoted '*Δpedestrians*', between stream populations observed in the last 100 updates of 100 replicate stochastic simulations over 10,000 simulation steps. Probability densities are shown separately for each value of γ (colour scale in panel A). Plots are for models 1 (A-C), model 2 (D-F) and model 3 (G-I). Values of *Δpedestrians* larger than zero indicate by how much at least two stream populations differ. For models 1 and 3, we observe non-zero *Δpedestrians* which suggests that stream populations can differ. In contrast, for model 2, values of *Δpedestrians* remain close to zero. This is expected from our stability analysis which suggests that at equilibrium all stream populations have the same value for this model. Differences between stream populations are due to stochastic fluctuations in simulations. Parameter values used are $\mu=0.62$ for models 1 and 2. For model 1, $\varepsilon=0.04$ and $\alpha=3$, *6* and *8*, respectively in (A-C). For model 2, $\varepsilon=0.01$ and $\alpha=1$, *3* and *6*, respectively in (D-F). For model 3, we use $\alpha=6$, $\mu=1.2$, $\delta=0.01$ and $\varepsilon=0.04$, *0.053*, and *0.1*, respectively in (G-I).



Supplementary figure S6: Results of probability integral transform applied to observed time intervals between events, Δt (i.e. time interval between separate pedestrians entering or leaving the crossing; see supplementary text for details). If the assumption underlying the Gillespie algorithm of exponentially distributed time intervals holds, the transformed time intervals should follow a uniform distribution between 0 and 1, indicated by dashed horizontal line in the plots. (A) shows the transformed Δt distribution obtained from selecting for each experimental run the parameters and model from the posterior distribution that produced a distribution closest to the uniform distribution (best-case scenario). For some experimental runs the match of the transformed Δt to a uniform distribution is much closer than the aggregated data shown. (B) shows the transformed Δt distribution obtained from selecting for each experimental run the parameters and model from the posterior distribution that produced a distribution furthest away from the uniform distribution (worst-case scenario). (C) shows the transformed Δt distribution obtained from sampling 1,000 times from the posterior distribution of models and parameters for each experimental run.